

University of Groningen

A panel data analysis of travel demand.

Meurs, Hendrik Jan

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1991

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Meurs, H. J. (1991). *A panel data analysis of travel demand*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

8. SUMMARY AND CONCLUSIONS

It is expected that travel demand will grow considerably in the next decade, with negative effects with regard to congestion, environment and safety. These developments are caused by a large number of demographic and socio-economic factors. This calls for new policies to bring about changes in behaviour. However, to achieve such aims will be a difficult task; behaviour appears to be relatively stable over time. One of the tasks of applied transportation research is to provide planners with insights into the contributing factors to mobility changes and the reason for apparent stability of travel behaviour.

Researchers have relied heavily on cross-sectional data and associated models in describing travel demand. This study deals with the potential contribution of panel data. Panel data provide us with measurements on a large number of households (the between dimension) over multiple points in time (the within dimension). This allows us to relate changes in mobility directly to changes in the contributing factors. By doing this, we control for unobserved heterogeneity; the permanent or time-invariant differences in mobility between households that cannot be directly attributed to their observed characteristics.

In chapter one it is argued that panel data analysis has a number of advantages in comparison with analysis based upon cross sections. One of the greatest advantages of panel data is the ability to check for misspecification due to left out variables. If omitted variables are correlated with the variables included in travel demand models, which is likely to be the case, the cross-sectional estimator will be biased. If the effects of these variables stay constant for a household over time, panel data analysis allows for elimination of these biases.

Another advantage of panel data is that dynamic models at the micro level can be estimated. The presence of short- and long-term effects implies that cross-sectional estimator will be biased. The common conjecture that cross-sectional analysis represents long-term effects is not valid in general; estimation of short- and long-term effects requires a dynamic model.

Furthermore, panel data allow estimation of models describing joint decision making with less restrictions needed to identify the direct and indirect effects than models based upon cross-sectional data.

In addition to the issues concerning model specification and appropriate representation of travel behaviour in the models, panel data allow the analysts to gain a detailed understanding of the reasons for stability in behaviour. Cross-sectional analysis attributes stability in behaviour to the fact that mobility determining factors change only slowly. For example, the spatial structure within which choices are made is relatively fixed over time.

However, there are other explanations for the apparent stability. If behaviour adjusts slowly to new circumstances, it may take time before changes in the mobility-determining factors will have full effects; long-term effects may be greater than short-term effects. Another explanation is that stability in decision is due to the fact that this decision is related to others which do not change; a single decision may change only if the set of factors affecting more than one decision changes. Finally, stability may be due to specific preferences and life styles. For example, some households are more mobile than others, and such differences may be particularly robust. To understand the mobility developments taking place and the reasons for the apparent stability, it is important to isolate and test these phenomena. Cross-sectional data are not rich enough to support this goal, since we only have data on differences between households.

In this study panel-data methods are used to estimate the effects of household characteristics on trip generation and car ownership. Comparisons of simple panel data models are made with estimates based on cross-sectional analysis. Subsequent analysis was carried out to test several hypothesis concerning the reasons for these differences. The Dutch Mobility Panel is used to estimate the models. Chapter two provides a description of this data set.

The starting point in the analysis is the fixed-effects model, described in chapter three. For each household a dummy variable is defined, in addition to the time-varying observed household characteristics. These dummies capture the effects of the time-invariant unobserved effects or unobserved heterogeneity. Estimation is performed by transforming the dependent and the independent variables in terms of deviations from their time-averages and subsequently applying Ordinary Least Squares. The major advantage of this model is that it controls for a correlation between the omitted variables and the included explanatory variables. This model was used to estimate coefficients describing the total number of trips made by households and the number of trips made by car and transit. A comparison of

fixed-effects
major diff

- the eff
- impact
- the cr
- reduce
- smaller
- effects
- transit
- Change
- Estima
- income
- to othe
- in total
- Income
- section
- license
- making

These resu
might over
on.

Next, the
heterogene
more mobi
heterogene
in conventi
making and
of the vari
not purely
unobserved
invariant u
rences with
wide range
imply that
These relat
included e:

appropriate repre-
sents analysts to gain
insight. Cross-sectional
analysis is determining
the effects within which

heterogeneity. If behaviour
changes in the
effects may be
heterogeneity in decision is
likely to not change; a
single more than one
preferences and life
circumstances, and such
heterogeneity developments
are important to isolate
heterogeneity enough to support
the model holds.

effects of household
comparisons of simple
cross-sectional analysis.
insights concerning the
model used to estimate the

described in chapter
in addition to the time-
measure the effects of
heterogeneity. Estimation is
variables in terms
using Ordinary Least
Squares for a correlation
model. This model
of trips made by
transit. A comparison of

fixed-effects coefficients with those from cross-sectional models revealed some major differences:

- the effects of increases in car ownership and licence holding have smaller impacts on the number of trips made by transit than expected on the basis of the cross-sectional model. Therefore, proper control for omitted variables reduces the effects of these characteristics. Also income effects appear to be smaller. Hence, cross-sectional methods might overstate the (negative) effects of future increases in income, license holding and car ownership on transit usage.
- Changes in income and license holding do not affect total trip generation. Estimates from the cross-sectional model were positive. Hence, the positive income effects for total trip generation in the cross-sectional model are due to other, unobserved effects. Changes in income will need lead to increases in total trip generation.
- Income effects are smaller in the fixed-effects model than in the cross-sectional model describing trip generation by car. In addition, changes in license holding and car ownership have less pronounced effects on trip making by car than expected from the cross-sectional model.

These results appear to indicate that cross-sectional analysis of trip generation might overstate the effects of observed household characteristics on trip generation.

Next, the contribution of the time-invariant unobserved effects (or unobserved heterogeneity) was examined; the phenomenon that some households remain more mobile than others over time. It was shown that significant unobserved heterogeneity is present in the data. Heterogeneity captures 30% of the residual in conventional models for total trip generation, 35% of the residual in car trip making and 60% in the transit usage equation. Therefore, a considerable amount of the variability in behaviour, after controlling for observed characteristics, is not purely random behaviour, rather it can be attributed to relatively stable unobserved differences among households. The high contribution of time-invariant unobserved effects in trip generation by transit may imply that preferences with respect to public transport are more idiosyncratic; households show a wide range of preferences with respect to transit. Also, specific constraints may imply that some households use transit very frequently, while others do not. These relatively stable preferences and constraints are not described well by the included explanatory variables. Such effects suggest that improvement of the

transit system will lead to substantial increases for some households, while others will continue to make very very trips.

Subsequent analysis aimed at testing a number of hypotheses generated to explain the differences between the cross-sectional and the fixed-effects model.

There are two possible observational explanations for these differences: panel attrition and measurement error.

Attrition is potentially a major problem of panel data. Households that have a higher probability of ceasing panel participation are not a random group. This may yield biased parameter estimates, explaining certain differences between the parameter estimates of the fixed-effects and the cross-sectional model. Ridder's study (1990b), described in chapter two, did not find a major impact of attrition on the regression coefficients associated with the household characteristics describing trip generation. However, the intercept of trip generation models is biased due to attrition.

Another explanation is the presence of measurement error. This error might not be higher than the error in cross-sectional surveys, but the fixed-effects model is more susceptible to bias resulting from such errors, providing an explanation for the lower effects of the household characteristics resulting from the fixed-effects model compared to the cross-sectional model. However, since we have panel data, we can test for measurement error and estimate the true coefficients. Tests, described in chapter three, show that measurement error does not contribute heavily to the differences between the parameters in the fixed-effects and the cross-sectional models.

The apparent minor contributions of attrition and measurement errors to the explanation of the differences leads to the conclusion that more substantive reasons are present for these differences. We explored two major substantive hypotheses: the presence of lagged effects implying that the differences are due to an underspecification of the dynamic character of decision making and, secondly, the presence of simultaneity.

First, consider the potential contribution of omitted dynamic effects in the static fixed-effects model. It is often argued that the fixed-effects model provides coefficients of transient effects, while coefficients in cross-sectional models

provide in
tics. When
For exam
the explan
On the oth
explanator
transforma
tions are r
tory varia
estimates

To examin
tion for th
ogeneity in
respect to
the first o
the genera
about the

The concl
the differ
describing
time-invar
generation
the short t
have some
the time-i
suggesting
term.

Next, we
explaining
models. In
exogenous
between t
failure to
generation
correlation

households, while
generated to explain
s model.

differences: panel

holds that have a
random group. This
ferences between the
al model. Ridder's
impact of attrition
old characteristics
eneration models is

this error might not
ed-effects model is
an explanation for
om the fixed-effects
nce we have panel
coefficients. Tests,
oes not contribute
ed-effects and the

ment errors to the
more substantive
major substantive
differences are due
sion making and,

effects in the static
ts model provides
s-sectional models

provide insights into the long-term effects of changes in household characteristics. Whether this is true depends upon the history of the independent variables. For example, if trip generation in previous periods affects current mobility and if the explanatory variables do not change substantially, this conjecture may hold. On the other hand, the fixed-effects model may provide short-term effects of the explanatory variables, if these regressors are uncorrelated over time after transformation in terms of deviations from their time-averages,. These assumptions are unlikely to hold, but differences in the autocorrelations of the explanatory variables before and after transformation may explain differences in the estimates if dynamic effects are present.

To examine the hypothesis that dynamic underspecification provides an explanation for the differences, requires a model that takes both lagged terms and heterogeneity into account. The estimation of such a model raises questions with respect to the initial conditions; what assumptions are necessary about the way the first observations are generated as a result of an unobserved past? We used the generalized method of moments, since this does not make any assumptions about the generation of the first observations.

The conclusion is that the dynamic misspecification explains only a small part of the differences between the fixed-effects models and the cross-sectional models describing trip generation. Most of the strong stability over time is explained by time-invariant unobserved characteristics. An interesting exception is trip generation by transit, for which the long term effects are about 18% higher than the short term effects. This implies that changes in household characteristics will have somewhat higher effects in the long term. It also implies that about 3% of the time-invariant residual variance can be explained by this phenomenon, suggesting that stability in trip making by transit is somewhat lower in the long term.

Next, we examined whether simultaneous equations models can contribute in explaining differences between the fixed-effects models and the cross-sectional models. In this study, we examined whether the treatment of car ownership as exogenous to trip generation is a tenable assumption. In particular the correlation between the unobserved effects is of concern. In chapter one it is argued that failure to take into account the correlation between the unobserved effects on trip generation and car ownership yields biased cross-sectional estimates. If the correlation between these unobservables is entirely attributable to the time-

invariant unobserved effects, the fixed-effects estimator will be unbiased, since this correlation is differenced out. Hence, the difference between the fixed-effects and the cross-sectional estimator may be due to different treatment of the correlation between the unobservables.

Another reason for considering such models is that they can provide us with a further qualification of the heterogeneity present in the data. If one studies a single decision which is related to other ones, omission of the inter-dependence may imply that heterogeneity in other decisions is attributed to the decision under study.

To examine these issues, we adopted a stepwise approach. First, models were estimated describing car ownership decisions within a linear framework (chapter five). This provided us with information about the presence of simultaneity. Obviously, this is a simplification; car ownership and mode usage decisions should be modelled within a joint discrete-continuous setting. Because the nonlinearities involved with these models poses some additional estimation issues, separate attention was given to discrete choice models describing car ownership decision making (chapter six). The results obtained were used in the joint model of car ownership and mode usage (chapter seven).

The linear simultaneous models were estimated with two assumptions about the effects of unobserved heterogeneity on decision making with respect to car ownership and mode usage. One set of models assumes that each decision is affected by separate, but correlated, unobserved time-invariant effects. In these models relatively high proportions of the total residual could be attributed to heterogeneity: 65% in the car ownership model, 55% and 50% for trip generation by car and transit, respectively. These random effects are significantly correlated, especially between the car ownership and car usage; about 25% of the heterogeneity can be attributed to common unobserved factors. The problem with this model is that it appears to be over-fitted. Therefore, another model was estimated with a one-factor structure. This more restrictive model uses one common latent variable to capture the common omitted variables affecting all decisions. This factor explained 66% of the residual for car ownership and 20% for car usage. For transit it explained only 5% of the total residual variance. Although the model appears to have less drawbacks in terms of over-fitting, the effects of car ownership on car usage were insignificant. This is implausible and

be unbiased, since between the fixed-treatment of the

n provide us with a ta. If one studies a he inter-dependence ted to the decision

First, models were framework (chapter ice of simultaneity. ode usage decisions etting. Because the dditional estimation dels describing car ed were used in the).

assumptions about the with respect to car at each decision is ant effects. In these uld be attributed to 0% for trip generati- ts are significantly sage; about 25% of actors. The problem , another model was ve model uses one riables affecting all ownership and 20% l residual variance. of over-fitting, the s is implausible and

may be due to the linear character of the model; it motivated a search for nonlinear specifications.

From the explorations, hypotheses were generated with respect to a further qualification of heterogeneity. It appears that a considerable amount of the unobserved heterogeneity in the car usage model is related to unobserved heterogeneity in car ownership. Transit usage appears to be much less related to car ownership and usage than expected. This may suggest that the stability in transit usage cannot be attributed to stability in car ownership and usage. To examine these hypotheses joint discrete-continuous models of car ownership and mobility were estimated.

The first stage of this analysis aimed at providing insights into the characteristics of heterogeneity in car ownership representing the phenomenon that households may have a high propensity to have a car, even in circumstances with increasing car costs and a relatively high quality public transport system. Results indicate that substantial unobserved heterogeneity is present with respect to car ownership. The parametrization of these time-invariant unobserved effects with a normal distribution may be untenable. The decision whether or not to own a car seems to be affected by heterogeneity with a J-shaped distribution, implying that a high proportion of the households exhibit relatively high propensities to own cars. Hence most households are likely to keep their car, even if the observed circumstances change to less car favoring circumstances. Only a small proportion of the households have a high probability to remain carless even if the observed household characteristics favor car ownership. The decision whether or not to own two cars is affected by unobserved time-invariant effects with a more symmetric distribution suggesting these households will respond more frequently to changing conditions affecting car ownership.

The cross-sectional model displayed a small correlation between the decision whether or not to own a car and the decision whether or not to own two cars; results from the one-factor model indicate that this is due to unobserved heterogeneity. However, the results also indicated that the assumption of independent decisions hardly biased the parameter estimates; so that the decisions concerning the first and the second car in the household can be considered to be independent.

The second stage of the analysis, presented in chapter seven, describes a joint model of car ownership, treated as discrete choices and mobility, defined in terms of number of trips. The estimated models take into account the interdependent nature of decision making with respect to car ownership and mobility.

From the analysis the conclusion may be drawn that car ownership and mobility decisions are correlated. These correlations can only be partially attributed to common exogenous variables affecting both decisions. It is important to take the correlations among the error terms into account as well. This affects a number of coefficients, especially those associated with income and accessibility. The direct effects of income changes are more positive for car usage and lower for transit usage; leading to the conclusion that income growth will be more favourable towards car usage and less for transit. However, the effects of improving the relative accessibility of transit with respect to car usage seem to be larger, implying that policies aimed at improving the position of transit will have more effects than expected on the basis of models which do not take these effects into account.

Another conclusion is that significant proportions of the unobserved time invariant effects in car trip generation models can be attributed to heterogeneity in car ownership decisions. Hence, the stability in car usage is significantly related to car ownership decisions. Hence, it will be difficult to affect car usage decisions without changing car ownership. Especially ownership of the second car might be subject to policies.

These findings confirm the findings obtained with the linear model: car ownership and trip making by car are related by the time-invariant unobserved effects. Hence, the differences between the cross-sectional and the fixed effects model describing trip generation by car can be explained to some extent by correlated random effects. Trip making by transit appears to be hardly related through these effects.

These findings have a number of consequences:

1. It may be difficult to affect mobility in terms of numbers of car trips without influencing car ownership decisions directly.
2. Single equation fixed-effects models provide researchers with coefficients of direct effects; parts of the differences between the fixed-effects and the

describes a joint
mobility, defined in
the interdepend-
ent mobility.

ship and mobility
ally attributed to
ortant to take the
ffects a number of
ibility. The direct
lower for transit
more favourable
of improving the
em to be larger,
it will have more
these effects into

erved time invari-
eterogeneity in car
ificantly related to
ar usage decisions
second car might

model: car owner-
unobserved effects.
xed effects model
tent by correlated
ated through these

umbers of car trips

s with coefficients
ed-effects and the

cross-sectional models can be attributed to the omission of indirect effects through car ownership.

3. Description of trip making by transit requires a dynamic model. There are some differences between short- and long-term effects for transit usage.

Summarizing, the major benefits of panel data are discussed and illustrated using simple models of trip generation. The benefits of panel data analysis are:

1. Such data allow us to examine whether the coefficients of travel demand models represent the effects of changes of household and individual circumstances on mobility.
2. They allow decomposition of a number of sources of variability and stability in behaviour.
3. Important sources of misspecification of conventional single equation models can be investigated, which can point at directions for new research.
4. Insights can be gained into the characteristics of the large unexplained variance of cross-sectional models.
5. Less restrictive assumptions need to be made about the structure of decision making and about the characteristics of the data.

Panel data analysis has only a short history in the field of travel demand research. Many issues and difficulties remain to be addressed. For example, the effects of changes in transport system characteristics, including time and cost, have not been addressed here; only recently have data become available with respect to these issues. One of the advantages of panel data is that variation in costs over time can be used to estimate cost elasticities. Combining such time variation with models derived from economic theory may yield new insights. This study attempted to describe the methods that can be used in such analyses.

Another important issue to pursue is the estimation of models capturing several decisions simultaneously, without the restrictions the nested logit model imposes. For example, the relationships between housing and labour markets and mobility decisions may be studied in more detail using panel data.

An issue of importance, but not addressed in this study is how to use the panel models for forecasting purposes. Cross-sectional models are used to forecast and

an issue is whether the models outlined in this study will improve the forecasting capabilities of transport planning models.

Finally, panel data can be useful in the analysis of relationships among decision making of household members. Changes in the number of employed persons within households may have large effects on the mobility characteristics of households.

This dissertation attempts to introduce a number of methods useful for analysis of panel data. We hope that future analysis will benefit from this work.